Article

Data Journalism as a Service: Digital Native Data Journalism Expertise and Product Development

Ester Appelgren [1,*] and Carl-Gustav Lindén [2]

[1] Department of Journalism, Södertörn University, 141 89 Huddinge, Sweden; E-Mail: ester.appelgren@sh.se
[2] Swedish School of Social Science, University of Helsinki, 00014 Helsinki, Finland; E-Mail: carl-gustav.linden@helsinki.fi

* Corresponding author

Abstract
The combined set of skills needed for producing data journalism (e.g., investigative journalism methods, programming, knowledge in statistics, data management, statistical reporting, and design) challenges the understanding of what competences a journalist needs and the boundaries for the tasks journalists perform. Scholars denote external actors with these types of knowledge as interlopers or actors at the periphery of journalism. In this study, we follow two Swedish digital native data journalism start-ups operating in the Nordics from when they were founded in 2012 to 2019. Although the start-ups have been successful in news journalism over the years and acted as drivers for change in Nordic news innovation, they also have a presence in sectors other than journalism. This qualitative case study, which is based on interviews over time with the start-up founders and a qualitative analysis of blog posts written by the employees at the two start-ups, tells a story of journalists working at the periphery of legacy media, at least temporarily forced to leave journalism behind yet successfully using journalistic thinking outside of journalistic contexts.

Issue
This article is part of the issue "Digital Native News Media: Trends and Challenges" edited by Ramón Salaverría (University of Navarra, Spain).

1. Introduction

Embedded in the terms 'digital native' or 'online native' is a precondition that there are (now) news sites without an analogue heritage. Harlow and Salaverría (2016) suggest that digital native news websites are not only innovative start-ups associated with new and alternative ways of producing and presenting journalism with differing degrees of political activism but also funded, or rather underfunded, by new business models. Furthermore, these initiatives are not limited to producing journalism, but also engage in the organization of cultural events, thus expanding on the business models for news production (García-Avilés, Carvajal-Prieto, Arias, & De Lara-González, 2019). According to Harlow and Salaverría (2016), digital native outlets tend to describe their use of technology

as different from mainstream media, for example, using data journalism techniques, such as creating databases and infographics and making documents publicly available. These techniques may not be different from what legacy media use, but they are nevertheless associated with being innovative. Developing skills in technology and self-promotion are important for entrepreneurial journalists (Cohen, 2012), a group that includes data journalists. Data journalists, in turn, are often regarded as forerunners of the journalism of the future (Knight, 2015). Loosen, Reimer, and De Silva-Schmidt (2017) argue that data journalism is commonly carried out by cross-disciplinary teams that have divided the labor into data analysis, visualization and writing. When professionals and non-professionals produce news together in this manner, they are engaging in what Belair-Gagnon

and Holton (2018) define as boundary maintenance (news production) around a boundary object (news). Researchers have described the professionals traditionally outside of journalism who are interacting at this boundary as interlopers or actors at the periphery of journalism (e.g., Belair-Gagnon & Holton, 2018; Cheruiyot, Baack, & Ferrer-Conill, 2019; Eldridge, 2018). According to Carlson (2016), these boundaries provide shared ways of understanding news work, and the metajournalistic discourse that arises from these collaborations may challenge how news is produced and consumed.

Previous research on news sites born online have focused on stand-alone sites with a direct audience (e.g., Bruno & Kleis Nielsen, 2012). In this study, we aim to explore how digital news initiatives that act as an intermediary between data and legacy media could also be part of the digital native ecosystem. Such news start-ups publish content and journalistic products, but not with the intention to attract an audience. Rather, the intention is similar to that of the larger non-profit news organizations around the world, such as ProPublica or the International Consortium of Investigative Journalists, which is to attract new collaborations with the media and through these collaborations republish the news to a wider audience (Konieczna & Powers, 2017).

This exploratory study revolves around a case featuring two such digital native news start-ups: Journalism ++ Stockholm and Journalism Robotics/Newsworthy. These companies share founders and employees and have overlapping aims of harnessing data at the periphery of traditional journalism. While Journalism ++ Stockholm mainly produces traditional data journalism on commission for legacy media, Journalism Robotics/Newsworthy focuses on automating parts of the data journalistic working process, i.e., the collection and refinement of public data, and provides analyses of data as short reports for both journalistic and non-journalistic purposes.

Hermida and Young (2019) assert that a key for future research is to gauge who matters in journalism based on qualities other than traditional ones (p. 99). Thus, the narrow focus on two individuals in this case study can be motivated by their central status as journalism innovators in Sweden and the Nordic countries. Furthermore, in his seminal work on news automation, Diakopoulos (2019) frames these two start-ups as groundbreaking in an international comparison as well.

Much of recent scholarship on innovation in journalism revolves around boundary work, where new actors enter journalism as peripheral actors (Powers & Zambrano, 2016; Usher, 2017) or engage with journalists as in the case of the entanglements of civic tech and data journalism (Cheruiyot et al., 2019). This qualitative case study is based on an analysis of blog posts and interviews over time with the start-up founders. The study contributes to the growing body of literature on peripheral actors in journalism by exploring digital natives who do not have a direct audience and apply a reversed perspective, thus presenting a story of journalists who work at the periphery of legacy media and use journalistic thinking outside of journalistic contexts.

In the following sections, we will briefly review the literature on digital news start-ups and relate data journalism to a subsection of the discussion on peripheral actors in journalism. We will then present the methods, material and motivating questions for the study, followed by our results. We conclude with a discussion and conclusions.

2. Literature Review

In this case study of two news start-ups, we focus on individuals who drive innovation and entrepreneurship in journalism from the periphery. Thus, the selected literature review is based on studies of entrepreneurial journalism, data journalism and peripheral actors, as well as factors of success in digital news start-ups.

2.1. The Rise and Success of Digital News Start-Ups

The literature on digital news start-ups revolves around start-ups producing journalism for a direct audience. In a report on such news start-ups in western Europe, Bruno and Kleis Nielsen (2012) found that online start-ups begin as journalistic and are driven by both professional and commercial ambitions. They tend to diversify their revenues by experimenting with activities that traditionally have not been associated with journalism, such as content syndication, e-commerce, advertorials, consultancy work, events planning or reader donations (p. 96). The scholarly discussion links this development of experimentation with new technologies and challenging of boundaries for what is considered journalism to the downsizing of the media industry. In this context, downsizing has had a positive effect by encouraging journalists to start their own businesses and propose their own journalistic innovations (Cohen, 2015; Raviola, 2019; Tenor, 2019). Cook and Sirkkunen (2015) divided digital native start-up sites into two categories: storytelling and service-orientation. Storytelling sites provide news to a direct audience, and service-oriented sites consist of online media actors that focus on elements of journalism rather than content creation. The revenue streams for service-oriented sites are not based on advertising models but rather on a wider strategy to sell products as part of a complex relationship with mainstream media (Cook & Sirkkunen, 2015, p. 73). Thus, previously unthreatened boundaries between journalistic and business-oriented functions are dissolving, often rhetorically motivated by survival and an industry crisis (Coddington, 2015). However, the change may not be that drastic. Naldi and Picard (2012) argue that entrepreneurs bring "formal myopia" into their new enterprise, which means their previous experiences and perceptions affect their assessments. Formal myopia, according to Naldi and Picard (2012, p. 77), influences initial thinking and choices of news start-up founders, thus limiting innovation and creating additional obsta-

cles to their success. However, Wagemans, Witschge, and Deuze (2016) argue that a strong belief in the core values of journalism, such as maintaining a professional identity, may not necessarily obstruct the development of new forms of entrepreneurial journalism; in fact, it can be the main selling point.

Some digital natives have been associated with counter information as a way to promote social change in societies facing democratic challenges and low confidence in the media (Harlow & Salaverría, 2016). There are also digital natives that focus on investigative journalism without distinct political or activist aims (Nee, 2013). Furthermore, there are non-profit digital news start-ups have emerged alongside market-driven online natives (e.g., in the US). Wagemans et al. (2016) found that French journalism start-ups are often inspired by US businesses that strive for fact-based, objective journalism. Similarly, for Nordic journalists, US journalism often serves as inspiration. There is a long history of the Americanization of European journalism, for example with norms, routines, and textual conventions previously being adapted to European national contexts (Broersma, 2019). However, Örnebring (2009) argues that, in a Nordic context, "Americanization is largely confined to news format, not news content" (p. 11).

Wagemans et al. (2016) note that most start-ups fail despite good intentions. One explanation for this could be that founders tend to primarily focus on developing their editorial model and spend little time to focus on revenues during the first few years (Naldi & Picard, 2012). In the case of journalistic start-ups, perceived success has been tied to the ability to offer quality content (Pekkala & Cook, 2012). To secure outreach, the start-ups also define success in terms of community, public service and survival rather than money. Therefore, Naldi and Picard (2012, p. 90) use a three-dimensional definition of initial success: 1) ability to attain original expectations, 2) capacity to change the business model, and 3) probability of survival (sustainability). Similarly, Bruno and Kleis Nielsen (2012) argue that a measure of success for digital news start-ups is mere survival. However, the ability to survive can be explained by the ability to deliver a distinct, quality product, operate with a lean organization, have diverse revenues, and target niche audiences that are poorly served by existing media (p. 6).

2.2. Data Journalism as a Form of Entrepreneurial Journalism

Cheruiyot et al. (2019) argue that data journalism is not a purely "journalistic" phenomenon because it is associated with peripheral actors who in various ways complement or expand the work journalists do (p. 2). For example, Nordic data journalists connect with each other through social media, such as the Facebook group "Datajournalistik," at conferences focused on data and journalism (Appelgren, 2016). In Europe, data journalism has been shaped by institutional factors, such as the

digitization of public records. These records are increasingly accessible via Application Programming Interface (API) mandated by the PSI Directive (Appelgren & Nygren, 2014) and the increasing supply of tools for collecting, analyzing, and visualizing data. Stalph and Borges-Rey (2018) assert that data journalism has become an asset within legacy organizations over the past decade. However, in this context, the nearby metaphor "robot journalism" has proved to have negative, almost existential, connotations (Lindén & Dierickx, 2019; Willis, 2020).

Based on previous research on boundary work in science, Carlson (2015) suggests a matrix to situate individual studies of boundary work in journalism. The model is based on studies of the separate lines between journalists and non-journalists. In a typical study of participants that move across Carlson's suggested categories of expansion, expulsion, and protection of autonomy, Baack (2018) found that data journalists and other professionals, such as interloping newcomers (e.g., civic technologists), engage in communities of practice, meaning that individuals rarely revolve around sustained and institutionalized engagement (p. 676).

While previous studies on entrepreneurial journalism have focused on pioneering forms of journalism in newsrooms, they have somewhat neglected individual pioneering journalists (Hepp & Loosen, 2019, p. 2). Similarly, even though data journalism is carried out in teams, there is little research on individual data journalists as pioneers. Pioneering journalists have been found to perceive themselves as forerunners who can act as intermediaries and bring together various competencies (Hepp & Loosen, 2019), and the ideal entrepreneurial journalist as embodied in the entrepreneurial discourse is an individual who creates new media products, disrupts traditional media organizations, and generates capital, new products, and even jobs for other journalists (Cohen, 2012, p. 524). In this context, there is an imperative to strive towards upgrading one's own skills, in particular concerning technology and self-promotion. Summarizing the scholarly research on data journalism, Loosen et al. (2017, p. 3) find that scholars are debating if and to what extent data journalism is actually a new reporting style, but the skills of the "technical journalist," such as programming and data journalism, are currently upgrading journalism (Bakker, 2014).

3. Background and Motivating Questions

In this section, we will briefly introduce the Nordic media market and provide the background on the two start-ups. We will then present the motivating questions for the case study.

The media ecosystem in Sweden, Norway, Finland and Denmark has been labelled the Nordic welfare state media model (Syvertsen, Mjøs, Enli, & Moe, 2014). It is characterized by strong public service and decentralized press, with a lot of local and regional newspapers (Hallin & Mancini, 2004). Social media platforms, and

Facebook in particular, have recently become the most frequent source of news for people in this region (Olsen, Solvoll, & Gran, 2018), but there are signs of a general trend away from dependence on platforms (Lindén, 2020). Legacy media is struggling with its business model, but the larger commercial media companies, such as Bonnier, Schibsted, Sanoma and Amedia, are successfully transitioning from mixed revenues—advertisement and subscriptions—to paid digital content (Villi et al., 2019).

The first start-up, Journalism ++ Stockholm, was founded by Swedish-Finnish journalist Jens Finnäs and Swedish journalist Peter Grensund in 2013 (Andén, 2013) in Stockholm. It was part of the European Journalism++ network with chapters in Paris, Berlin, Stockholm, Porto, Amsterdam, and Cologne. Later, the Nordic agency added three new employees. The larger European Journalism++ network was founded in 2011 by Nicolas Kayser-Bril, Pierre Romera, and Anne-Lise Bouyer (Kayser-Bril, 2017). The most famous work from this network was the award-winning project "The Migrants' Files". The European network Journalism++ coordinated this groundbreaking project, which aimed to measure the number of people who died while trying to reach Europe, between 2013 and 2016 (The Migrants' Files, 2014). The network shrank to only two chapters in 2017: Stockholm and Porto.

The second start-up, Journalism Robotics, was first described in 2016 as the product Newsworthy, and later presented as a separate company, Journalism Robotics. Newsworthy is described as a "machine that will find news in data" (Journalism Robotics/Newsworthy, 2020).

J++ Stockholm has been active in the global open-source movement, sharing software code and practices on platforms such as GitHub. Furthermore, it attended key data journalism conferences and seminars, not just in Sweden and Nordic countries but also internationally, such as the European gathering Dataharvest and the US NICAR and Computation + Journalism Symposium conferences. With the emergence of artificial intelligence in journalism, the news media is becoming increasingly dependent on technical expertise (Beckett, 2019), and helping newsrooms understand and make use of technology to produce various forms of data-driven journalism is at the center of the J++ Stockholm and Journalism Robotics/Newsworthy business model. In order to explore how the start-ups negotiated their area of expertise in data journalism over time as they develop the company and their skills, we use two motivational questions: 1) "What are the aims and target groups, i.e., customers, for the two news start-ups?"; and 2) "How do the founders of the news start-ups define their success and failure in navigating the boundaries between journalism and technology?"

4. Methods and Material

This case study is based on a qualitative content analysis of blog posts from the J++ Stockholm and Journalism Robotics/Newsworthy websites and a series of semi-structured interviews with the main founder, Jens Finnäs, and a cofounder, Måns Magnusson, of two news start-ups, Journalism ++ Stockholm and Journalism Robotics/Newsworthy.

Case study research is an empirical process that investigates a contemporary phenomenon, and it is iterative in nature (Yin, 2014). Bryman (2008) argues that it is not always possible to discern the case study type until after a detailed investigation. However, because the case selected for this study is an example of the more general case of a peripheral actor in journalism in the Nordics, we believe that it can be viewed as what Yin (2014) referred to as a representative case.

The interviews with the two founders were conducted in Stockholm and Helsinki, respectively, on November 24, 2016, August 9, 2017, and March 15, November 8, and November 11, 2019. The interviews lasted around 40–50 min each. The interviews were transcribed and analyzed using a thematic and inductive qualitative text analysis approach, where careful reading of transcripts has formed the motivating questions that structure the article. This method has been applied in a flexible way, taking into account the nuances and diversity of the responses. The researchers are positioned as active in the research process; themes do not just emerge (Braun & Clarke, 2006). However, the text analysis is grounded in the subjective meaning of human action, thus preserving the interviewees' subjective point of view (Fereday & Muir-Cochrane, 2006).

To ensure validity, we showed the informants the results of the interview study and corrected any misunderstandings. The corrections were mainly about the order of events in time, the narrative of leaving journalism behind, and the labelling of the companies and products as start-ups. Because the material mainly relates to specific circumstances regarding a small number of individuals, we applied an idiographic approach to the analysis (Bryman, 2008).

For the content analysis, we collected the blog posts available on both companies' web pages. In total, Journalism ++ Stockholm and Journalism Robotics/Newsworthy created 23 blog posts between June 2, 2015 and October 16, 2019 (18 and 5, respectively). The majority of Journalism ++ Stockholm's posts were written by two authors: nine by Jens Finnäs, seven by Leo Wallentin, and the remainder by other staff members. The Journalism Robotics/Newsworthy posts do not include bylines. When coding, the researcher looked for the process of narration in the material, i.e., how the information was offered, withheld or delayed (Gillespie, 2006), or how the company positioned and described itself and its development. We decided to use a qualitative content analysis of company blog posts, where one of the researchers first coded for specific themes in the textual content (data journalism, investigative journalism, technology, non-journalistic practices, success, failure, companies and organizations) and then looked for how these

themes were narrated when promoting the company's aims and activities.

We regard the Journalism ++ Stockholm blog posts as "physical traces of how organizations represent and account for themselves" (Coffey, 2014, p. 367), and the purpose has been to gather information about "sense-making practices" (McKee, 2003, p. 52). The material was read several times and coded (Bryman, 2008). After the initial coding, we found that the blog posts illustrated company development by accounting for the company's activities over time. The results from the content analysis follow the motivating questions and are organized in terms of the intention and language for whom the content was created and for what purpose, as well as the cultural resources the stories draw on and what they aim to accomplish with the stories (Riessman, 1993). All quotes are translated from Swedish.

5. Results

In this section, we will present the results from the content analysis of blog posts and the interviews with the founders of the two digital journalism start-ups.

5.1. Results from the Content Analysis of Blog Posts

An overview of the two blogs forming the basis for the content analysis shows that they utilize different styles. The Journalism ++ Stockholm blog has a personal tone and uses colloquial language with a rich flora of data journalism terms that might be new to non-data journalists. Instead, the Journalism Robotics/Newsworthy blog primarily addresses non-journalists and promotes Newsworthy as a product. However, there are similarities between the blogs in terms of rhetoric. Both blogs frequently use questions at the beginning of texts and paragraphs. Statistics and results from analyses are presented with dialogical language, referring to the reader as "you" and the writer together with the reader as "we," thus guiding the reader in a personal style through complex descriptions of data journalism methods. Both blogs contain images and graphics primarily for the purpose of illustrating relationships found in data.

The first motivational question of our study, as presented earlier, relates to the aims and target groups. We found that the blog posts accounts for company aims by describing development related to data journalism competence and skills. Target groups are mentioned in examples of collaborations with media companies or non-journalistic organizations. Several narratives illustrate the founders' experiences of becoming more skilled at data journalism and developing their company ideas. We therefore begin this section by presenting three quotes that are particularly illustrative of how the company positions itself using personal examples and nestled promotional stories. The first quote provides a historical account of how Journalism ++ Stockholm first was established:

Four years ago, Jens Finnäs and Peter Grensund received an offer they couldn't refuse: an opportunity to join the Journalism ++ network. They founded Journalism ++ Stockholm and became part of an international network of agencies dedicated to data-driven journalism and newsroom innovation. (Finnäs, 2017)

This quote illustrates how the two journalists entered the international network of data journalists as a successful strategic move. The blog post continues by praising the competence in the international Journalism ++ network, but its main purpose is to mention that key members of the international Journalism ++ network are leaving. This is carefully explained so that it does not seem as if this will do any harm to those remaining in the Journalism ++ network. The blog post then explains that a smaller network will lead to a new exciting era with opportunities to tighten collaborations and become even stronger at producing data journalism.

In the second quote, the presented narrative not only describes the journalistic experience, but functions as a suggestion and an argument for why 'robot journalism,' the main business model of Journalism Robotics/Newsworthy, is valuable to news organizations:

I began my career as a journalist at the local newspaper Nya Åland. It is not uncommon for local news reporters to feel a bit like a robot. Rewriting police telefax statements as news items about speeding was not very intellectually stimulating. Why couldn't a news robot have done this job? (Finnäs, 2015b)

This quote is followed by calculations of costs that are possible to cut in the newsroom, thus clearly targeting journalists in charge of media houses. The quote aims to motivate senior journalists now in management positions to invest in news automation technology to eliminate mundane tasks. However, in several posts, the start-up founders make it clear that they are not interested in discussing job cuts and replacing journalists with software, but rather the possibilities presented by new technology. The presented experience the founders acquired from having a full career, from the mundane and simple tasks at the junior level as in the quote above, to the present, where expertise in programming makes it possible to reduce such tasks, still implicitly indicates that the mundane experience has been valuable.

The third quote focuses on experience, and is found in the beginning of a Journalism Robotics/Newsworthy blog post on how to create graphs:

I have been the leader of courses and workshops on how to visualize and tell stories with data intended for journalists and Public Relations professionals for more than ten years. What should you bear in mind when you create a graph? What is the secret behind good data visualization? Here, I have tried

to reduce everything I have learned in three steps. (Newsworthy, 2019a)

The post appears to be about presenting useful information in graphs, but it implicitly promotes both the Newsworthy product, which provides analytics as data, and the competence of the product's creators. The promotional examples given in the post are primarily non-journalistic and use personal pronouns: "Sales have gone down since *we* changed *our* CEO, or housing prices in *our* municipality are up" (Newsworthy, 2019a). Furthermore, the signal to the reader is that they might not be competent enough in creating graphs. This is visible in other blog posts as well, for example by referring to the audience as having no time for developing skills or journalists notoriously being bad at math. While the quote above is rather humble, the style quickly changes in the blog post. The insights gained from ten years of teaching visualization are summarized as what most journalists do wrong, rather than what journalists can do right, and the author's tone can be interpreted as annoyed or even superior:

> This is by far the most important step. Let me clarify: The most I-M-P-O-R-T-A-N-T. (Newsworthy, 2019a)

The second motivational question of our study is related to critical factors for success. The blog posts discuss the company's success by listing successful projects. On rare occasions, successes are described explicitly, but not by the authors themselves. For example, the following is a press release that indicates success through the quote of a new employee:

> This is very exciting! Newsworthy and J++ [Stockholm] are among Sweden's most innovative media players. (Newsworthy, 2019b)

The post continues with quotes from the founders about how happy they are that the group will now be able to carry out even more skilled data journalistic tasks. Other posts include comments from previous course attendees praising the teaching and what they have learned by attending Journalism ++ Stockholm courses.

The general finding from the analysis of blog posts, however, is that the focus is not on explicit success but rather on presenting the methodology behind successful projects. There is a strong emphasis on the difficulties each project faced. Thus, success is primarily narrated as being transparent with learning new things, as captured in the following quote:

> One of our internal goals at J++ [Stockholm] is for us to work with aspects that are new to us in all our projects. There has to be something innovative in the genre or something that we ourselves have not yet mastered. (Wallentin, 2015)

The tension between technology and journalism and the applied journalistic thinking to non-journalistic areas was visible in the blog posts. We found that substantial effort was made to put future 'robot journalism' in a positive light, listing strengths such as efficiency, endurance and ability to avoid human error. The 'robots' are said to be better at judging quality, for example when interpreting statistics.

In a blog post aimed at promoting an early version of Journalism Newsworthy by describing how 'robot journalism' works, a timeframe is used to impress journalists. Two hours after a public agency releases data, a news automation tool will have analyzed the published data and created 311 reports. In other words, the impossible is now possible. The blog post continues by describing how this process is effortless, as results "drop in from all over the country." This process is also enjoyable:

> We have not only written 1,244 reports (with only two people at the wheel), but we have also created a machine that can create 1,244 reports while we are having a cup of coffee. (Finnäs, 2018)

In this context, mentioning the coffee break signals that it is possible to get large amounts of work done while taking a break. Again, the impossible is now possible. While the blog posts also state that additional human competence is needed to carry out journalistic tasks, and that humans are shaping technological processes, for example that humans have written the news automation software, 'robots' are nevertheless presented as living beings with feelings:

> Carrying out repetitive tasks are one of robot journalism's primary strengths. For the robot, it does not matter if it will get an assignment to write one police report or a thousand. It is just as happy. (Finnäs, 2015b)

Gradually, blog posts increasingly contain non-journalistic examples. Results, or rather methodology behind projects in collaboration with public agencies and unions, shape the message to appeal more to professional but possibly non-journalistic target audiences. Yet the reader is constantly reminded that the products and expertise can be used for journalistic purposes:

> Newsworthy started as an initiative for helping local journalists find news in data. Now, we are obviously also helping elected members in the Teachers' Union understand their world with data....In our news service Newsworthy, we try to build a machine that can answer these questions *en masse* for different types of data. If you just keep the questions at the back of your head the next time you present numbers to your boss, you will have a much better shot at making your voice heard. (Newsworthy, 2019c)

5.2. Results from the Qualitative Interviews with the Start-Up Founders

This section summarizes the interviews with the two founders and is structured around the two motivational questions. The first motivational question, as presented above, concerns the aims and target groups for the two news start-ups. In a report written by Jens Finnäs, he defines the start-up's aim as follows:

The ambition is to be the leading data journalism freelance agency in Scandinavia, an actor that newsrooms across the Nordic countries can turn to for assistance in data-driven reporting. (Finnäs, 2015a)

In essence, this is the definition of a service provider, and this aim is also reflected in the company's blog posts through their emphasis on describing competence and data journalism methodology. Finnäs explained the logic of the two different corporate structures as two commercial operations with slightly different focuses. According to Finnäs (personal communication, November 11, 2019), Journalism ++ Stockholm is a consultancy firm that helps media companies with various journalistic projects and also trains reporters in data journalism, while Journalism Robotics/Newsworthy has a broader user group, including people in public relations and communication. According to Finnäs, the infrastructure for local content can easily be turned into press releases, personalized emails, or PowerPoint presentations as well as articles.

While the aim of Journalism ++ Stockholm and the reasons for starting the network were explicit in the analyzed blog posts, the interviews revealed that the idea behind Journalism Robotics/Newsworthy emerged from a friendship with statistician Måns Magnusson. Initially, Finnäs had been a freelance data journalist and explored how open public data sources could be used for journalism. Finnäs was friends with a statistics PhD student, Måns Magnusson, who was focusing on machine learning. They were both inspired by US statistician Nate Silver and his start-up FiveThirtyEight, which uses statistical analysis to tell compelling stories.

Before receiving his PhD degree, Magnusson worked as a statistician for the Swedish National Council for Crime Prevention, a popular data source among Swedish data journalists. Magnusson and Finnäs started collaborating professionally, one focusing on hardcore data mining and the other on searching for interesting data that could be used and visualized in journalistic projects. In 2016, their collaboration resulted in Journalism Robotics, which developed news automation services under the brand name Newsworthy. Magnusson (personal communication, November 8, 2019) described the service as "computational storyfinding, going through large amounts of data with statistical methods, finding a needle in the haystack." Jens Finnäs asserts that the start-up's journalistic logic of working with structured data is appealing to more than just journalists:

Newsworthy adjusts created texts to local contexts and patterns to create a sophisticated software pipeline that autonomously generates texts with very little human effort in terms of adjusting content. We are building the infrastructure to create local content…providing 300 different local news angles instead of one national angle. (Finnäs, personal communication, November 11, 2019)

Apart from producing news content on commission for larger Nordic news organizations, the company also applied for funding from non-profit organizations and foundations, although projects funded in this manner generally were terminated once the funding was gone. One example of this was a project to develop a data mining system for a large newspaper in Tampere, Finland, that failed to gain acceptance in the newsroom. The failure of technology adoptions in newsrooms is a common thread in research (Wagemans & Witschge, 2019).

When the Google-funded Digital News Initiative (DNI) started in 2015, Finnäs and Magnusson decided to apply for funding. They received funding for an idea of news automation based on data sets from Statistics Sweden, a government agency that produces and publishes official statistics. In interviews, Finnäs describes this moment as a defining and critical success factor for the company, and the idea developed into the start-up news service Newsworthy. Overall, Google has been a crucial supporter of data journalism and media innovation (Fanta, 2018; Lindén, 2020). In 2016, Magnusson and Finnäs were involved in an incubator program managed by the Swedish innovation agency Vinnova. This project was plagued by communication problems and what was perceived as less-than-helpful advice: "Don't get me started on that," Magnusson said afterwards (personal communication, November 8, 2019). The main problem was that the funding from this program could not be used to develop their current business model; rather, it could only be used to develop some kind of tangible product. Finnäs and Magnusson struggled to find a solution that Vinnova was willing to fund, and they were asked to "think outside the box" and "be brave." Ultimately, they received no funding.

Reflecting on the project funded by DNI, Finnäs (personal communication, November 11, 2019) believes that they were doing things quite differently today compared to their original pitch. Newsworthy's main function was supposed to be to provide local newsrooms with unique content, and more specifically localized news snippets. Because of the lack of commercial interest among Swedish newsrooms, though, they decided to broaden its services to other customers. Newsworthy landed new customers outside of journalism, such as the Swedish teachers' union, *Lärarförbundet*.

The second motivating question of this study, as presented above, is related to perceived success and failure.

A key part of the perceived success has been to educate journalists and take part in conferences. According

to Finnäs, "Through all these workshops, we have seen that the data skills and Excel skills are very, very low" (personal communication, November 11, 2019). Finnäs has also held a key role in the Nordic Data Journalism Conference and built the repository for entries to NODA Awards as part of attracting new collaborations: "Conferences have been totally invaluable for finding coworkers and customers. For me personally, everything can be traced back to some conference" (Finnäs, personal communication, November 11, 2019).

Finnäs accounts the failures to a lack of interest for the Journalism Robotics/Newsworthy product. The perceived non-response from newsrooms to their offer was disappointing, especially since Finnäs is confident that the tool could help overcome structural obstacles in editorial routines for an industry in which both money and time are scarce.

We found in the interviews that the founders primarily told stories about designing and building the systems, but they somewhat neglected their strategies for addressing target groups to be successful. Finnäs is self-critical when he reflects on how to get people already struggling with information overflow to listen and accept that even more new information could be useful:

> The challenge is not technical, more like organizational. Looking at municipalities, we need to know who will gain and in what way? What are their roles, are they bureaucrats or elected? What topics are interesting and relevant and where do they get their information today? (Finnäs, personal communication, November 11, 2019)

Magnusson is more optimistic, asserting that the low level of interest is only temporary and Newsworthy will be able to focus on journalistic services in the future:

> Since the amount of available data is quickly growing, interesting stories will drown in an unmanageable mass of information. With the help of machine-learning tools such as Newsworthy, journalists can build systems that go through massive amounts of data to find both anomalies and large societal trends. (Magnusson, personal communication, November 8, 2019)

6. Conclusions

Journalism ++ Stockholm and Journalism Robotics/Newsworthy are well-known in the Nordic media market and viewed as pioneers driving change in Nordic news innovation. As such, they are examples of non-traditional actors that matter in journalism (Hermida & Young, 2019). Yet, in this case study, we have found that the company founders are now strategically moving towards non-journalistic customers. With this study, we aim to expand the digital native news start-up research by exploring two digital natives without a direct audience

that are working at the periphery of the legacy media, successfully managing to apply journalistic thinking outside of journalistic contexts.

Two motivational questions guided us through our exploration of the digital journalism start-ups' journey. The first motivating question concerned aims and target groups for the two start-ups.

Journalism ++ Stockholm and Journalism Robotics/Newsworthy initially were examples of what Cook and Sirkkunen (2015) denoted as service-oriented digital native sites; they focused on elements of data journalism and had a strategy to sell products to legacy media. In line with García-Avilés et al. (2019) and Bruno and Kleis Nielsen (2012), we also found that they offer alternative products, such as organizing events, and they diversify their revenues with consultancy work and content syndication. To face economic realities, Journalism ++ Stockholm and Journalism Robotics/Newsworthy have gradually become more flexible in terms of their focus on data journalism as a service for legacy media and successfully attracted a few non-journalistic customers.

The founders have not put much effort into researching the needs of either their journalistic or non-journalistic target groups. However, interviews and company blog posts reveal that the founders are very knowledgeable about what can be done with data in a journalistic setting and how to create quality content, an important presumption for news start-up success (Pekkala & Cook, 2012). There is also evidence that they have mastered how to collect and structure data, for example by offering transparent explanations of methodologies, but we did not find explicit offers for attracting customers to the Newsworthy product. The absence of a customer needs analysis beyond the news production process might be an example of what Naldi and Picard (2012) denote as "formal myopia," thus restricting innovation and development beyond the competences and experiences that the founders brought with them into the companies.

The second motivational question revolves around how the companies define their success and failure in navigating the boundaries between journalism and technology.

We argue that the decision to continue with the operations, despite the failure of the larger European J++ network, shows that mere survival, as suggested by Bruno and Kleis Nielsen (2012), was a factor of success. However, training journalists with computational skills has probably been the most important success factor for J++ Stockholm, establishing their competence niche in data journalism. Furthermore, throughout the company's existence, it has successfully produced data journalism for several newsrooms across Scandinavia on commission. This is in line with how Wagemans et al. (2016) argues that core values of journalism and maintaining the professional identity can be part of the main selling point for new forms of entrepreneurial journalism. In general, however, we found that Journalism ++

Stockholm expresses success in a quite negative manner. However, explaining gained insights in blog posts as overcoming failures, a transparent way of expressing skill development may be well-suited to attract data journalistic projects on commission. Traditional media companies were willing to pay the founders for their freelance data journalism content and data journalism training courses in a traditional buyer–seller relationship.

However, we found that the Nordic media companies were reluctant to invest in journalistic products, such as Newsworthy. We argue that the strong emphasis on the contested metaphor 'robot journalism' in the promotional discourse could have been an obstacle. Robot journalism is associated with job loss and thus brings negative connotations to newsroom staff (Lindén & Dierickx, 2019), furthermore, robots are portrayed as more than just automation of mundane tasks, and may thus appear threatening. Therefore, even though journalism forms the basis of the two start-ups, the Journalism Robotics/Newsworthy product, with its emphasis on robotics, may have appeared as challenging, presenting contradictory definitions of the profession (Eldridge, 2018). When trying to sell the product as technology-enhancing journalism, the founders move from the buyer-seller relationship between journalists, as seen when selling data journalism projects on commission, to outsiders, i.e., interloping actors (Belair-Gagnon & Holton, 2018). With the narratives of efficiency and automation, a competing narrative is provided. Reluctance could thus be an example of how legacy media protects autonomy by fending off non-journalists seeking control or who aim to shape journalism (Carlson, 2015).

To conclude, our case study shows that the two start-ups engage heavily in developing technology skills and use the blogs for self-promotion in a manner typical of entrepreneurial journalists (Cohen, 2012). They found a successful narrative to attract data journalistic work on commission acting as journalists, yet become outsiders when selling their product Newsworthy to legacy media. They master the advantage of what Carlson (2016) describes as the metajournalistic discourse regarding cooperation between journalists and technologists by transparently describing failures and methodological challenges in data journalism in their promotional blog posts. However, they have yet to learn how to identify customer needs and communicate what the product can offer in a manner that is appealing to both journalistic and non-journalistic customers. Nevertheless, they attract non-journalistic customers.

Future research could investigate if this type of promotional communication needs to be formulated differently in order to sell the product in non-journalistic contexts and at the same time appeal to the professional logic of journalism. Perhaps the core issue in selling the product Newsworthy is not the fault of the two start-ups, but rather highlights the shortcomings of legacy media to understand why and how they also need to cooperate with peripheral actors.

Acknowledgments

Conflict of Interests

The authors declare no conflict of interests.

References

Andén, A. (2013, January 23). Vad är affärsnyttan med datajournalistik, Jens Finnäs? [What is the business value of data journalism, Jens Finnäs?]. *Medievärlden*. Retrieved from https://www.dagensmedia.se/nyheter/medievarldens-arkiv/vad-ar-affarsnyttan-med-datajournalistik-jens-finnas

Appelgren, E. (2016). Data journalists using Facebook. *Nordicom Review*, *37*(1), 156–169.

Appelgren, E., & Nygren, G. (2014). Data journalism in Sweden: Introducing new methods and genres of journalism into "old" organizations. *Digital Journalism*, *2*(3), 394–405.

Baack, S. (2018). Practically engaged: The entanglements between data journalism and civic tech. *Digital Journalism*, *6*(6), 673–692.

Bakker, P. (2014). Mr. Gates returns: Curation, community management and other new roles for journalists. *Journalism Studies*, *15*(5), 596–606.

Beckett, C. (2019). *New powers, new responsibilities: A global survey of journalism and artificial intelligence*. London: The London School of Economics and Political Science.

Belair-Gagnon, V., & Holton, A. E. (2018). Strangers to the game? Interlopers, intralopers, and shifting news production. *Media and Communication*, *6*(4), 70–78.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101.

Broersma, M. (2019). Americanization or the rhetoric of modernity: How European journalism adapted US norms, practices and conventions. In K. Arnold, P. Preston, & S. Kinnebrock (Eds.), *The Handbook of European communication history* (pp. 403–419). New York, NY: Wiley.

Bruno, N., & Nielsen, R. (2012). *Survival is success: Journalistic online start-ups in western Europe*. Oxford: Reuters Institute for the Study of Journalism.

Bryman, A. (2008). *Social science research methods* (3rd ed.). Oxford: Oxford University Press.

Carlson, M. (2015). Introduction: The many boundaries of journalism. In S. C. Lewis & M. Carlson (Eds.), *Boundaries of journalism* (pp. 1–18). New York, NY: Routledge.

Carlson, M. (2016). Metajournalistic discourse and the meanings of journalism: Definitional control, boundary work, and legitimation. *Communication Theory*, *26*(4), 349–368.

Cheruiyot, D., Baack, S., & Ferrer-Conill, R. (2019). Data journalism beyond legacy media: The case of African and European civic technology organizations. *Digital Journalism*, *7*(9), 1–15.

Coddington, M. (2015). The wall becomes a curtain: Revisiting journalism's news–business boundary. In S. C. Lewis & M. Carlson (Eds.), *Boundaries of journalism* (pp. 67–82). New York, NY: Routledge.

Coffey, A. (2014). Analysing documents. In U. Flick (Ed.), *The SAGE handbook of qualitative data analysis* (pp. 367–380). London: SAGE.

Cohen, N. S. (2012). Entrepreneurial journalism and the precarious state of media work. *South Atlantic Quarterly*, *114*(3), 513–533.

Cook, C., & Sirkkunen, E. (2015). What's in a niche? Exploring the business model of online journalism. *Journal of Media Business Studies*, *10*(4), 63–82.

Diakopoulos, N. (2019). *Automating the news: How algorithms are rewriting the media*. Cambridge, MA: Harvard University Press.

Eldridge, S. A., II. (2018). *Online journalism from the periphery: Interloper media and the journalistic field*. New York, NY: Routledge.

Fanta, A. (2018). The publisher's patron: How Google's News Initiative is re-defining journalism. *European Journalism Observatory*. Retrieved from https://en.ejo.ch/digital-news/the-publishers-patron

Fereday, J., & Muir-Cochrane, E. (2006). Demonstrating rigor using thematic analysis: A hybrid approach of inductive and deductive coding and theme development. *International Journal of Qualitative Methods*, *5*(1), 80–92.

Finnäs, J. (2015a). *Report from NICAR 2015*. Helsinki: Finnmedia. Retrieved from http://www.medialiitto.fi/files/3132/Jens_Finnas_matka-apuraha_NICAR-konferenssiin_2015.pdf

Finnäs, J. (2015b, June 29). Därför behöver vi fler robotjournalister [That is why we need more robot jouranlists]. Retreived from jplusplus.org/sv/blog/darfor-behover-vi-fler-robotjournalister

Finnäs, J. (2017, April 11). Journalism++ enters a new era. Retrieved from jplusplus.org/sv/blog/journalism-enters-a-new-era

Finnäs, J. (2018, January 19). 1 244 lokala nyheter på 2 personer och 24 timmar [1244 local news reports created by 2 persons in 24 hours]. Retrieved from jplusplus.org/sv/blog/1244-lokala-nyheter-pa-2-personer-och-24-timmar

García-Avilés, J. A., Carvajal-Prieto, M., Arias, F., & De Lara-González, A. (2019). Journalists' views on innovating in the newsroom: Proposing a model of the diffusion of innovations in media outlets. *The Journal of Media Innovations*, *5*(1), 1–16.

Gillespie, M. (2006). Analysing media texts. In M. Gillespie & J. Toynbee (Eds.), *Narrative analysis* (pp. 79–117). Maidenhead: Open University Press.

Hallin, D. C., & Mancini, P. (2004). *Comparing media systems: Three models of media and politics*. Cambridge: Cambridge University Press.

Harlow, S., & Salaverría, R. (2016). Regenerating journalism: Exploring the "alternativeness" and "digitalness" of online-native media in Latin America. *Digital Journalism*, *4*(8), 1001–1019.

Hepp, A., & Loosen, W. (2019). Pioneer journalism: Conceptualizing the role of pioneer journalists and pioneer communities in the organizational re-figuration of journalism. *Journalism*. https://doi.org/10.1177/1464884919829277

Hermida, A., & Young, M. L. (2019). From peripheral to integral? A digital-born journalism not for profit in a time of crises. *Media and Communication*, *7*(4), 92–102.

Journalism Robotics/Newsworthy. (2020). Statistik innehåller nyheter: Vi hittar dem åt dig. [Statistics contain news: We find them for you]. *Newsworthy*. Retrieved from https://www.newsworthy.se/sv

Kayser-Bril, N. (2017). Things learned in 5 years of journalism++. *Nkb*. Retrieved from https://blog.nkb.fr/things-learned-jpp

Knight, M. (2015). Data journalism in the UK: A preliminary analysis of form and content. *Journal of Media practice*, *16*(1), 55–72.

Konieczna, M., & Powers, E. (2017). What can nonprofit journalists actually do for democracy? *Journalism Studies*, *18*(12), 1542–1558.

Lindén, T. C. G., & Dierickx, L. (2019). Robot journalism: The damage done by a metaphor. *Unmediated: Journal of Politics and Communication*, *2*, 152–155.

Lindén, T. C. G. (2020). *Silicon Valley och makten över medierna* [Silicon Valley and the power over media]. Gothenburg: Nordicom.

Loosen, W., Reimer, J., & De Silva-Schmidt, F. (2017). Data-driven reporting: An on-going (r) evolution? An analysis of projects nominated for the Data Journalism Awards 2013–2016. *Journalism*. https://doi.org/10.1177/1464884917735691

McKee, A. (2003). *Textual analysis: A beginner's guide*. London: SAGE.

Naldi, L., & Picard, R. G. (2012). "Let's start an online news site": Opportunities, resources, strategy, and formational myopia in startups. *Journal of Media Business Studies*, *9*(4), 69–97.

Nee, R. C. (2013). Creative destruction: An exploratory study of how digitally native news nonprofits are innovating online journalism practices. *International Journal on Media Management*, *15*(1), 3–22.

Newsworthy. (2019a, June 4). Hur gör man ett bra diagram? [How to create a graph?]. Retrieved

from https://medium.com/@contact_76790/hur-g%C3%B6r-man-ett-bra-diagram-e1cc69da9d9a

Newsworthy. (2019b, June 4). Newsworthy rekryterar från BBC [Newsworthy recruits from the BBC]. Retrieved from https://medium.com/@contact_76790/newsworthy-rekryterar-från-bbc-c2399e7b23fc

Newsworthy. (2019c, October 16). En ranking [A ranking]. Retrieved from https://medium.com/@contact_76790/en-ranking-290-powerpoint-presentationer-383becdb38d1

Olsen, R., Solvoll, M., & Gran, A.-B. (2018). *Digitalisering av lokal mediebruk* [Digitization of local media use]. Oslo: BI Centre for Creative Industries.

Örnebring, H. (2009). *Comparative European journalism: The state of current research*. Oxford: University of Oxford.

Pekkala, P., & Cook, C. E. (2012). Sustaining journalistic entrepreneurship. In E. Sirkkunen & C. Cook (Eds.), *Chasing sustainability on the net: International research on 69 journalistic pure players and their business models* (pp. 110–135). Tampere: Tampere Research Centre for Journalism, Media and Communication.

Powers, M., & Zambrano, S. V. (2016). Explaining the formation of online news start-ups in France and the United States: A field analysis. *Journal of Communication*, *66*(5), 857–877.

Raviola, E. (2019). Just like any other business or a special case? Framing excess in a Swedish newspaper group. In B. Czarniawska & O. Löfgren (Eds.), *Overwhelmed by overflows?* (pp. 96–110). Lund: Lund University Press.

Riessman, C. K. (1993). *Narrative analysis* (Vol. 30). Newbury Park, CA: SAGE.

Stalph, F., & Borges-Rey, E. (2018). Data journalism sustainability: An outlook on the future of data-driven reporting. *Digital Journalism*, *6*(8), 1078–1089.

Syvertsen, T., Mjøs, O. J., Enli, G. S., & Moe, H. (2014). *The media welfare state: Nordic media in the digital era*. Ann Arbor, MI: University of Michigan Press.

Tenor, C. (2019). Logic of an effectuating hyperlocal. *Nordicom Review*, *40*(s2), 129–145.

The Migrants' Files. (2014). *The migrants' files*. Retrieved from http://www.themigrantsfiles.com

Usher, N. (2017). Venture-backed news start-ups and the field of journalism: Challenges, changes, and consistencies. *Digital Journalism*, *5*(9), 1116–1133.

Villi, M., Grönlund, M., Linden, C. G., Lehtisaari, K., Mierzejewska, B., Picard, R. G., & Röpnack, A. (2019). "They're a little bit squeezed in the middle": Strategic challenges for innovation in US metropolitan newspaper organisations. *Journal of Media Business Studies*. https://doi.org/10.1080/16522354.2019.1630099

Wagemans, A., Witschge, T., & Deuze, M. (2016). Ideology as resource in entrepreneurial journalism: The French online news startup Mediapart. *Journalism Practice*, *10*(2), 160–177.

Willis, H. (2020). Journalism on autopilot: The upside and downside of computer-generated stories. *Quill Magazine*. Retrieved from https://www.quillmag.com/2020/01/14/journalism-on-autopilot-the-upside-and-downside-of-computer-generated-stories

Wagemans, A., & Witschge, T. (2019). Examining innovation as process: Action research in journalism studies. *Convergence*, *25*(2), 209–224.

Wallentin, L. (2015, June 15). En faktagranskningsmetod för datajournalistik [A factchecking method for data journalism]. Retrieved from jplusplus.org/sv/blog/en-faktagranskningsmetod-for-datajournalistik

Yin, R. (2014). *Case study research design and methods* (5th ed.). Thousand Oaks, CA: SAGE.

About the Authors

Ester Appelgren is an Associate Professor of Journalism and the Head of School of Natural Sciences, Environmental Studies and Technology at Södertörn University in Sweden. She as a PhD in Media Technology from the Royal Institute of Technology, Sweden. Her research interests focus on data journalism and digital integrity. During 2011–2014, Ester managed an action research-oriented project on data journalism, funded by the Swedish Innovation Agency (Vinnova).

Carl-Gustav Lindén is Associate Professor at the University of Helsinki and a Researcher in the intersection between journalism, media and technology with a focus on innovation and business models. He has led a large Finnish research project developing Natural-Language Generation systems for media (www.immersiveautomation.com) and is currently involved in creating new Artificial Intelligence tools for newsrooms working in lesser resourced languages (www.EMBEDDIA.eu).